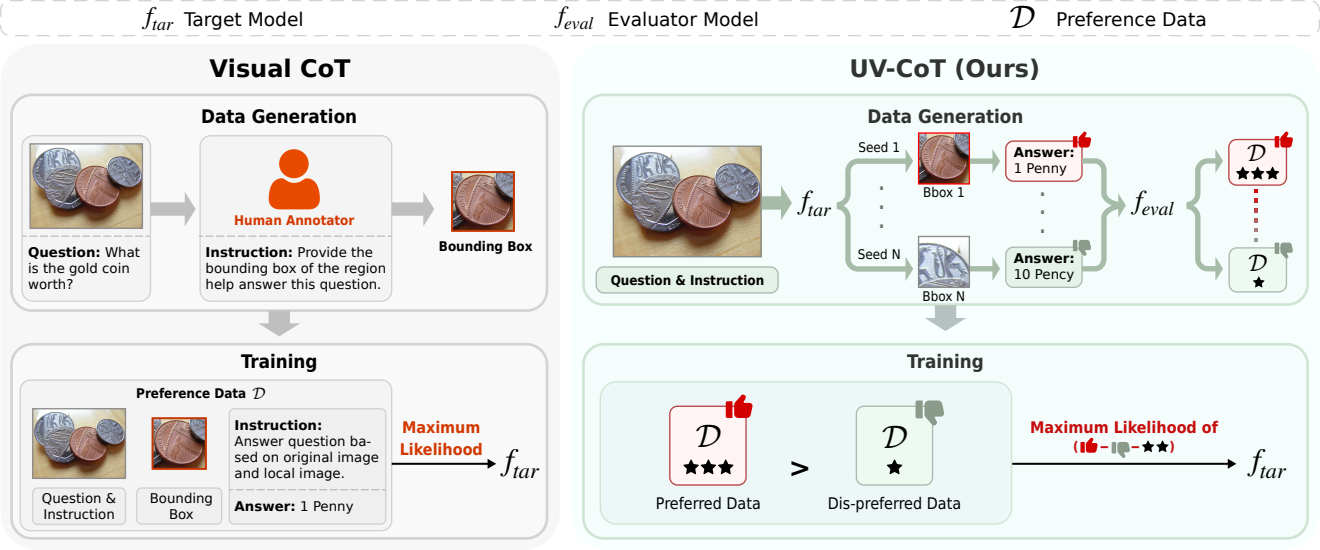
f_{tar} Target Model
f_{eval} Evaluator Model
$\mathcal{D}$ Preference Data
Visual CoT
Data Generation
Question: What is the gold coin worth?
Human Annotator
Instruction: Provide the bounding box of the region help answer this question.
Bounding Box
Training
Preference Data $\mathcal{D}$
Question & Instruction
Bounding Box
Instruction: Answer question based on original image and local image.
Answer: 1 Penny
Maximum Likelihood
f_{tar}
UV-CoT (Ours)
Data Generation
Question & Instruction
f_{tar}
Seed 1
Bbox 1
Answer: 1 Penny
Seed N
Bbox N
Answer: 10 Pency
f_{eval}
$\mathcal{D}$
$\mathcal{D}$
Training
$\mathcal{D}$
Preferred Data
>
$\mathcal{D}$
Dis-preferred Data
Maximum Likelihood of
f_{tar}